

# Southeast Information Node (SEIN)

## DRAFT STRATEGIC PLAN

2006 to 2009

---

### **I. Background/Purpose of the Plan**

During the first five years of existence, the Southern Appalachian Information Node (SAIN) has built a solid and respected information infrastructure. It has established a core group of collaborating partners who have worked with more than 100 agencies and organizations to identify and link more than 250 data sets, and crystallized new connections and connectivity among data generators and data users. Through data gathering and mapping efforts, it has leveraged external resources and helped solve specific practical, on-the-ground problems.

During its second generation, SEIN will develop a broader contiguous geographic focus in the Southeast. This is reflected in the name change from the Southern Appalachian Information Node (SAIN) to the Southeast Information Node (SEIN). SEIN seeks to identify the significant bioinformatics needs of the region; substantially advance the development of tools, content, and applications to meet those needs; and build a hub of unparalleled excellence and utility. This plan is designed to guide these efforts and serve as a strategic touchstone for directing and measuring progress.

### **II. SEIN's Mission**

Our mission statement describes SEIN's reason for existing:

*"Making the region's biodiversity information  
available for decision making."*

SEIN is part of a growing national and international network of biological information hubs designed to bring diverse biological data sources together to aid decision making. SEIN, which is part of the National Biological

Information Infrastructure (NBII), provides the information infrastructure for the Southeastern United States. Using innovative information and technology tools, it serves diverse decision makers, stakeholders, and communities within the region, as well as nationally and internationally.

### **III. Vision**

When fully developed, SEIN will be a major repository and gateway to high quality biodiversity and ecosystem information used to support areas such as resource management, sustainable development, research, and for education and outreach. SEIN will use and promote cutting edge information technologies, data processes and standards, and collaborative organizational relationships among data generators and data users so that barriers to the discovery, access, and dissemination of biodiversity data are dissolved.

SEIN will be widely recognized as a:

- Source for downloadable biodiversity information
  - Datasets resulting from biodiversity studies
  - Map layers representing species distribution, habitat, and other environmental elements
  - Data and information products from an array of biodiversity data sources
- Source for tools and applications for visualizing and analyzing biodiversity data
- Source for locating biodiversity resources by
  - Contributing to the catalog of web resources
  - Promoting development of metadata
  - Linking to metadata clearinghouses
  - Linking to biodiversity weighted search tools
- Partner with communities of biodiversity interest
  - Hosting data and information products
  - Supporting collaboration resources for the purpose of identifying information needs, data gaps, and developing community standards

### **IV. The Region Defined**

The NBII has established regional nodes for focusing and providing services within a particular geographic area of the country, with a strategic goal of covering the entire nation. Nodes have been delineated based on characteristics such as ecological uniformity, shared biological issues, and the availability of high quality biological and information science capabilities,

data, and expertise within the region. The Southern Appalachian Information Node (SAIN) was established in 2000 with the states of Kentucky, Tennessee, Alabama, and Mississippi comprising its general service area. Although many of SAIN's initiatives in first five years have been focused on East Tennessee and southern Appalachia, it has become apparent that the ecological uniformity and shared biological issues extend throughout the broader southeast. The discussion of redefining the geographic coverage of the node was included in the development of this strategic plan.

With this strategic plan, the states of Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Alabama, and Mississippi will comprise the general service region for SEIN. However, recognizing that the biodiversity and ecological issues of interest do not stop at state borders, they will continue to be addressed within the context of the biogeographical and physiographical regions originating within these states.

## **V. Priorities**

The SEIN general service region is one of the most biologically diverse areas within the United States and contains some of the fastest growing population centers. The combination of these two factors makes it imperative that SEIN address cross-cutting issues when addressing priorities.

SEIN intends to identify information to assist decision makers by collaborating with communities of knowledge to find information that already exists but is not currently available on the web.

Cross-cutting issues include

- Invasive species: informing and enabling prediction, early detection, and rapid response.
- Threatened and endangered species: collaboration for information collection with subject experts, and making appropriate levels of data accessible for both individual species and assemblages.
- Fire: Providing access to abundant but disparate research and data on fire as an ecological factor that affects not only regional ecosystems, but human lives and property as well.

In fulfilling SEIN's Vision the following have been selected as initial focus areas:

- Mountainous Ecosystems – Defined as that portion of the Appalachian range that extends from Pennsylvania's Alleghenies south to Alabama's Red Mountain. SEIN will focus primarily on the southernmost area of this region, but will ensure broad coverage through collaborations with the NBII Mid-Atlantic Information Node. The focus will be on hotspots, protected wilderness areas, and unique ecosystems.
- Forest Ecosystems – The USDA Forest Service (USFS) and state forestry agencies have expended tremendous money and effort on managing these ecosystems that are a significant portion of land within the SEIN region. SEIN intends to focus on identified information gaps and will work with the expertise of these agencies to increase information exchange.
- Major Water Ecosystems that Transect the Region – In particular, the Tennessee River and its tributaries, and the Lower Mississippi River and its tributaries. SEIN intends to focus on enabling access to biodiversity information that will contribute to decisions made in regard to water quality, quantity, and access.
- Urban /Wildlands Interfaces – Transition areas for urban, sub-urban, and industrial development with wildlands. SEIN intends to focus on the land use conversion dynamics and environmental implications of both human-wildlife interactions and conflicts, as well as impacts of species habitat loss and/or fragmentation.
- Agricultural Ecosystems – The SEIN region has multiple small-scale agri-ecosystems collectively having a tremendous impact on the biological resources of the region.

## **VI. Stakeholders**

- Tier 1: elected, appointed, and organizational decision makers; resource managers, program implementers.
- Tier 2: scientists, including citizen scientists, doing applied research.
- Tier 3: educators and students.

## VII. Assumptions About the Next Three Years

1. SAIN has successfully established a recognized presence in the region (including a website, an initial infrastructure, a suite of projects, etc.). The regional bioinformatics potential may not yet be fully realized. Without demonstrated outcomes in the next three to five years, SEIN may become increasingly vulnerable and realization of its goals may be obstructed.
2. Because resource management, conservation, and agriculture are critical to the future of the Southeast, there will continue to be a pressing need for reliable, high quality biological information that is at once molecular, temporal, and spatial. Also vital to the process is the accessibility of that information to the public in general, and to policy makers, resource managers, researchers and technical experts, citizen scientists, and educators, more specifically.
3. Security and environmental issues will become increasingly enmeshed as the public's awareness of human and environmental hazards increases.
4. Like other developing bioinformatics nodes, SEIN will continue to face ongoing strategic and operational "tugs and pulls," among them the dual challenges of developing a broad base of content while developing content and services pertinent to specific locales or problems, investment in infrastructure as opposed to investment in projects, and the cultivation of new partnerships while providing services to existing entities. In the next generation of its growth, SEIN will have to evolve from a collection of miscellaneous opportunistic projects to an enterprise that is more focused and deliberate. It must identify the significant needs in the region and focus resources on them. It must achieve a better balance of "breadth" and "depth" to demonstrate bioinformatics leadership within the region.
5. SAIN will continue to be a USGS-coordinated partnership of interested federal, state, and local agencies; educational institutions; and private and non-governmental organizations.
6. SAIN must increase the amount of content through greater data sharing with federal and state agencies and with organizations in the private and civil sectors.
7. Initial base funding for SEIN will continue to come primarily from USGS, but is likely to either stay the same or decline over the next three years. Between 2006 and 2009, any increases in funding will probably have to come from sources outside of the NBII base. NBII funds can be expected

to cover key operational costs, but resources for specific, unique projects must increasingly come from partners, customers, and others. While no partner has a lock on, or entitlement to funding indefinitely, organizations that cultivate new funds should be able to access and use those funds.

8. SEIN will contribute data openly and freely to facilitate data analysis and knowledge development. However, contributors will be assured respect for ownership of data and given appropriate citation.
9. The public is an increasingly sophisticated body of information users. There are increasing levels of impatience with slow webs, portals, and servers. Sites that do not maintain technological sophistication run the risk of becoming irrelevant. Computing will continue to advance toward more distributed systems.

## **VIII. Strategic Goals**

Between 2006 and 2009, SEIN will focus the greatest percentage of its financial, human, and organizational resources in regional bioinformatics on the activities, initiatives, projects, and services that demonstrate the fullest possible efficacy. SEIN's highest overall programmatic priority is to achieve tangible programmatic success in ways that demonstrate high value to diverse stakeholders, data users, data providers, and communities. To accomplish this, SEIN will focus on the following goals:

1. Substantially increase the amount of content relative to the SEIN priority focus areas.
2. Demonstrate a "Proof of Concept."
3. Advance SEIN's infrastructure and technologies.
4. Achieve long-term financial sustainability.
5. Expand to all states in the Southeast region.
6. Increase outreach and collection of user feedback.

## IX. Strategies

Following are the broad strategies SEIN will use to achieve its strategic goals. These are organized by the six strategic goals from Section VIII above.

1. **Create Greater Breadth.** Substantially increase the amount of “baseline” content, the number of data contributors, and the number of data users.  
Strategies:
  - a. Assemble a top tier “Advisory Committee” of representative regional stakeholders to (a) identify and prioritize the critical biodiversity/ecological issues of the region; (b) develop strategies for working with high priority science issues in the region; and (c) develop a prioritized set of data sets and databases.
  - b. Conduct an assessment and prioritization of the available data sets now held by federal and state agencies as well as others held in the private or civic sectors; i.e., research collections, resource collections, and reference collections.
  - c. Invest in and leverage the imaging/data mining research of partner universities to improve discovery of web-based information resources.
  - d. Create additional virtual and actual forums (portal hosted communities of knowledge, information bazaars, workshops, convened meetings, etc.) for the exchange of ideas among data users, data generators, and others interested in bioinformatics.
  - e. Create specific Memorandum of Understanding (MOUs), Memorandum of Agreement (MOAs), or letters of agreement regarding the development of infrastructure and the sharing of information.
  - f. Implement a competitive grant program to fund high priority data rescue/digitization efforts.
2. **Demonstrate a Proof of Concept.** Demonstrate that SAIN can (a) bring diverse biological data sets together; (b) apply specialized analytical, mining, or visualization tools; and (c) contribute to practical, on-the-ground solutions to real-world issues and problems.

- a. Meet as a core group to fully describe what a reasonable “Proof of Concept” project will entail, examine the kind of project that might best lend itself to the earliest possible proof, and create a working document to be circulated among those with biological and bioinformatics interest in the region. Projects to be considered for a prospective proof of concept project include these criteria:
  - i. Ability to actually impact real-life decision making, especially decisions with economic consequences
  - ii. Potential to bring multiple partners to the table
  - iii. Regulatory or legal drivers
  - iv. Fundability and the capacity to leverage other resources
  - v. Solid knowledge base
  - vi. Capacity for development of a media profile
  - vii. Regional impact

Develop at least one (and not more than three) possible project that could demonstrate the node’s fullest capability for addressing specific locale-based problems. Candidates could be projects that build on existing work, or new projects such as an inventory and capability for modeling the biota of the Tennessee River, a corridor mapping of the Appalachian Trail, or a predictive model of invasive species.

- b. As part of the project selection process, conduct an inventory and assessment of the available data sets.
- c. Meet with knowledgeable policy makers and resource managers to determine a problem to which SAIN can address its bioinformatics capabilities.
- d. Work with resource managers and other prospective partners to prepare a specific listing of data sets and metadata needed for the project and create specific procedures for data sharing.

3. **Advance SEIN's Technology and Infrastructure.** Improve SEIN's web design, search engines, and long-term infrastructure for maintaining content.
  - a. Establish a Technology and Infrastructure Working Group and Content/Knowledge Management Working Group to provide focused and consistent day-to-day management and operation of the SEIN core infrastructure.
  - b. Develop an information architecture for the SAIN core infrastructure to document core capabilities, standards, and services (data hosting, GIS, metadata development, training, etc.).
  - c. Establish a mechanism for user input and feedback on web design and content.
  - d. Establish mechanisms to ensure SAIN infrastructure is compatible with NBII network-wide infrastructure activities and standards.
4. **Achieve Financial Sustainability.** Maintain and increase base funding from USGS while developing new sources of funds to support specific projects and initiatives that meet specific needs in the region. As current USGS projects complete, redeploy USGS funds to improve operations and infrastructure development.
  - a. Consult with all project leaders and mutually determine effective end dates for current projects.
  - b. Collaborate with other NBII Nodes for cost-sharing projects.
  - c. Review existing operations and infrastructure and develop a technical and business plan for needed improvements and expansions.
  - d. Seek new projects and new sources of project funding. As the current projects supported solely by USGS conclude, seek new projects that bring resources with them. Where USGS-funded projects are to continue, begin creating financial matches.
  - e. Partner with universities and other bioinformatics initiatives (such as ATBI) to develop NSF (and other) grant proposals to support bioinformatics research initiatives.

- f. Partner with the DOE Oak Ridge National Laboratory and the Southern Appalachian Cooperative Ecosystems Studies Unit to develop National Ecological Observation Network (NEON) grant proposals (i.e., implement proof of concept application of the Corridor Mapping Application) as SEIN progresses.
- 5. **Expand SEIN's geographic activities to all Southeastern states.** Funds permitting, ensure that there is some type of SEIN presence in each of the region's states and bioregions.
  - a. Meet with federal and state agencies as well as researchers and citizen science groups in each state to articulate the biodiversity informatics challenges, data sets available, and needs and priorities in each of the states.
  - b. Meet with policy makers in the newly added states to explore future funding opportunities that would allow for SAIN's expansion into those areas.
- 6. **Increase outreach and collection of user feedback.** Developing strategies for outreach and collecting user feedback are essential to the growth and development of the SEIN.
  - a. SEIN will maintain core support for outreach activities which will include targeting exhibits to regional conferences, working with partner public relations activities, working with the NBII Coalition and developing articles and white papers for NBII publications.
  - b. User feedback will be solicited regularly to improve the usability and quality of the SEIN.

## **X. Stakeholders**

SEIN's stakeholders can best be represented as the collective communities within the region who generate biological information and/or need access to biological information and the IT resources to make this information useful. These stakeholders include:

- a. Policy makers faced with impending decisions
- b. Scientists conducting applied research
- c. Natural resource managers
- d. Heritage organizations
- e. Conservation organizations

- f. Schools and other educational institutions
- g. Others with specialized biological data interests or information needs

SEIN will continue to seek innovative and effective ways of interacting with, listening to, and involving diverse stakeholder groups by:

- a. Offering training programs and workshops that promote informatics standards such as Open GIS, cataloguing, metadata, and others.
- b. Participate through presentations and exhibits at major regional conferences such as the annual meetings of the Association of Southeastern Biologists, Southern Appalachian Man and Biosphere, and Discover Life in America All-Taxa Biodiversity Inventory.

## **XI. Measures of Success**

SEIN will track and evaluate its performance using a variety of metrics and reports, as follows:

### **Metrics**

- 1. Outcome stories: specific reports or projects and activities that demonstrate on-the-ground impacts (similar to the uses for decision making that occurred around Eastern Brook Trout and Oriental Bittersweet data).
- 2. Number of partners and data contributors.
- 3. Website traffic.
- 4. Completion of deliverables.
- 5. Presentation of papers by SAIN partners and staff.
- 6. Increase in storage and data.
- 7. Volume of content.

## **XII. Governance**

SEIN has multiple levels of leadership. The “Core” group consists of staff members, paid and volunteer, who have the immediate responsibility for developing, maintaining, and expanding SEIN’s infrastructure. “Partners” consist of data contributors and users, persons or organizations working on SEIN projects, and others who are also undertaking applied natural bioinformatics research and development.

From time to time, all partners will be invited to participate in meetings, workshops, or working groups that are designed to solve specific problems, develop or expand projects, consider pending decisions, and provide recommendations on specific issues.

The general operating principle is to make well-informed, highly inclusive, and thoughtfully considered decisions that will be most likely to achieve consensus. However, recognizing that consensus may not always be feasible for every decision, it is understood that the Node Manager has the final responsibility for making and implementing decisions.

Chartered working groups will be established to provide consistent, operational support and coordination for SEIN core functions. Staffed core functions are:

- a. Content/Knowledge Management
- b. SEIN Webmaster
- c. Partnership Development and Outreach
- d. Infrastructure Management

In addition to ongoing user feedback, the NBII Steering Committee and the SEIN Science Advisory Committee will provide analysis of SEIN’s adherence to this strategic plan and its success at addressing regional needs.